GLOSSOLOGIA 9-10 (1990-1991) 81-94

GREEK INFLECTION: A COMPUTATIONAL ANALYSIS IN A MACHINE TRANSLATION SYSTEM (EUROTRA)

ANGELA RALLI

In this paper, I report an attempt to analyse and generate computationally the inflectional system of Greek within the framework of the multilingual machine translation project known as EUROTRA. My work is based on the assumption that inflection constitutes a morphological process and assignment of morphosyntactic features to inflected items may be handled by using feature bundle representations and feature-passing operations. It is shown that the morphological process of inflection is particularly important to the computational treatment of languages like Greek with a rich inflectional system.

0. Introduction

A major issue in the study of morphologically complex words is the location of inflectional morphology. Anderson (1982) and Aronoff (1990) defend the hypothesis that inflectional processes are triggered by syntactic information and must, therefore, take place somewhere within a postlexical component[1]. This point of view differs from a strong lexicalist position (cf. Jensen 1990) according to which the lexicon generates inflected forms independently of syntactic requirements. My position, in this paper, is that inflection should be performed in an autonomous morphological component. I follow the lexical theory of generative morphology which accommodates a full range of inflected forms in Modern Greek (henceforth simply Greek). I argue that a well developed set of features and a principled method of feature-passing may set the frame within which theoretical and computational investigations of inflection might proceed. In this approach, the morphological entities involved in every inflectional process of word formation (i.e., stems and inflectional endings) are formulated in terms of feature bundles and inflectional structures appear as feature bundle representations. Inflection, then, becomes a morphological process of feature - matching and feature - passing between feature bundles in feature bundle representations.

The main aim of the paper is the presentation of the computational analysis of the Greek inflection. To this aim, I use the unification - based formalism adopted by EUROTRA, the multilingual machine translation project[2]. The EUROTRA model (henceforth simply ET) consists of several representational levels, both in analysis and generation, that is, a morphological level (EMS), a surface syntax level indicating order of constituents (ECS), a relational syntax level (ERS), and an interface structure (IS) which is a deep

---

1. In Bauer (1988) there is a detailed account of the criteria usually used for defining inflection as a syntactic process. He shows that all these criteria do not belong to a universal list and are open to serious objections.
2. More specifically the version of the years 1989-1990.

syntactic dependency representation containing governor - argument - modifier structures and featurised morphosyntactic and semantic information. Each level consists of a dictionary, a generator and a translator. The dictionary captures the relevant information,expressed by means of features (attribute/value pairs), for each lexical unit. The generator reflects information concerning grammatical structures and translators link contiguous levels.

The paper is mainly divided into two parts. The major characteristics of Greek inflection are presented in the first part and a number of theoretical issues with respect to the grammatical status of inflection are discussed next. Inflected items are considered to costitute morphologically complex words containing a stem and an inflectional ending. The second part of the paper is devoted to the computational treatment of inflection. All the evidence in this paper has been taken from Greek. Nevertheless, the inflectional systems of other European languages may be treated in a similar manner.

1. Greek inflection: theoretical remarks

1.1 General characteristics

Let me begin with the relevant details of a particularly rich system such as Greek inflection. The main feature characterising inflected words is grammatical category (noun, verb, adjective, etc.). Depending on the content of this information, several other morphosyntactic features mark the words. For example, case (nominative, genitive, accusative and vocative), number (singular and plural), gender (masculine, feminine, neuter), inflection class (i.e., declensions) and, for computational reasons, person (first, second, third) mark all nouns. Most of these features are shared by adjectives. As for verbs, the range of morhosyntactic features is rich and complex (e.g., mood, voice, aspect, tense, person, number, inflection class).

Inflection in Greek is carried out through affixation. Inflectional affixes are generally added to bound morphemes, i.e., stems, in order to form fully inflected words[3]. For example, a stem like 'kip -' «garden» gives rise to eight paradigmatic types (only seven of them are distinct in form) depending on the addition of the appropriate inflectional endings:

(1)	Singular	Plural[4]
Nominative	kipos	kipi
Genitive	kipu	kipon
Accusative	kipo (n)	kipus
Vocative	kipe	kipi

3. In the paper, the term «affix» will alternate with the term «ending», both widely used in the literature.
4. Greek examples here are transcribed phonologically according to the characters of the International Phonetic Alphabet. It should be noted though that the transcription used by EUROTRA-Greece is different and has been invented for the purposes of processing texts. Although I have chosen to ignore stress (EUROTRA - Greece does the same), it is worth saying that stress placement and stress shift phenomena constitute a case of phonology - morphology interaction (cf. Malikouti-Drachman and Drachman 1988, Touratzidis and Ralli 1992).

According to traditional grammatical descriptions, nouns are distributed into three and verbs into two inflection classes (term used to cover declensions for nouns and conjugations for verbs). With some exceptions, adjectives generally share the same inflection classes with nouns[5]. However, the old declensions and conjugations that these words are referred to do not have a real classification value because they do not reflect the actual division in inflectional paradigms. According to the shape of their inflectional endings, verbs, nouns and adjectives are distributed into several distinct paradigms, that is, members of the same paradigm show the same inflectional ending for any given proper morphosyntactic combination. The use of a special diacritic feature marking both stem and ending in order to ensure the right matching between them has already been proposed by Ralli (1986) as a replacement for the use of old declensions and conjugations. My interpretation of this diacritic feature states that membership of a given inflection class is indicated by a feature, the specific value of which triggers a particular inflectional ending for a particular stem bearing the same feature. Distinct values of this feature assume the classificatory role of distinct inflection classes and their number depends on the number of different paradigms. In this paper, I refer to the diacritic feature of inflection class with the abbreviation 'ic' and I use Arabic numerals for the specific classes:

(2) $kip_{ic:1}$ $os_{ic:1}$ «garden, nominative singular»

I say, for example, that for masculine nouns, a stem like 'kip-' belongs to inflection class [1], another stem such as 'maxiti-' (with its allomorph 'maxit-') «fighter» is inflected according to the paradigm of inflection class [2], etc. (for a detailed account of Greek inflection, cf. Ralli 1988).

Stems sharing the same 'ic' value belong to the same paradigm, that is, to the same set of inflectional forms. Replacement of old declensions and conjugations by a special feature has no effect on the system of other features. As a matter of fact, Meillet (1934) and Ernout (1953) had already observed that inflection classes are not direct carriers of morphosyntactic properties. For these authors, inflection classes have a classificatory role and exist as independent parts of the grammar. The same claim was further explored by Aronoff (1990) who pointed out that the classification system of inflection classes is motivated by an inherent human desire for order.

From the examples given above, it is obvious that in Greek inflection we do not find distinct affixes for every feature. Inflectional affixes are, therefore, portmanteau morphemes. Moreover, for each paradigm of nouns and adjectives there is a certain amount of syncretism. As Joseph and Philippaki-Warburton (1986) point out, syncretism is often, but not always, disambiguated syntactically by the use of different articles. For example, with the masculine noun 'patera' «father» and the feminine noun 'mitera' «mother» we find the use of different articles:

(3) Nominative / Vocative, singular: i, _ mitera
Genitive, singular: tu patera
Accusative, singular: ton patera, ti mitera

---

5. For example, the adjective 'evγenis/evγenis/evγenes' «noble» is an exception.

1.2 The theory of morphology: some basic points

Morphology deals with the structure of words and the ways in which this structure reflects their relation to other words. Words are composed of minimal signs called morphemes which combine a linguistic meaning (or a grammatical function) with a phonological representation. Analysing a word means identifying the distinctive meaningful elements within it. On the other hand, generating a word implies to reconstitute its message by combining its smaller units and their associated meanings.

Such a treatment presupposes the existence of constraints on how morphemes can be combined. Well - formulated constraints of this sort can be found in all morphological theories; however, in this paper, I have chosen to follow the guidelines of lexical generative theory of the 80' s because, in my opinion, it constitutes the most recent well-elaborated framework of all morphological studies.

The meaning of a complex word requires the representation of certain relations of scope among its components. This is realised by assigning to words an analysis that implies a kind of internal hierarchical structure. Within the framework of lexical generative grammar, the word internal structure is analysed and generated by allowing the application of a small number of context - free word formation rules.

Words, stems and affixes appear as terminal elements in tree-representations of word structures. In all languages, however, these morphological units are not always distributed in the same manner. For example, English nouns belonging to the class of words are combined with the inflectional plural marker '-s'. This does not hold for Greek where only stems can be combined with inflectional affixes.

1.3 Inflection as a morphological process

According to Jensen (1990: 63), a morphological process can be defined as one that simultaneously changes the form of a morpheme, usually a stem, and adds an element of meaning to that morpheme. Following this observation, Greek inflection can be defined as a morphological process because by joining an ending to a stem, the new form that occurs is followed by an addition of a certain amount of information to the basic semantic interpretation of the stem. What is unique to inflection though is that inflection provides forms of lexemes, while another morphological process such as derivation, for example, provides new lexemes. The question, then, is the following: is this difference between inflection and derivation so radical that it calls for a different treatment in distinct components of the grammar? In other words, should inflection be taken into account by syntactic rules while derivation is left in morphology? The syntactic use of inflectional information drove Anderson (1982: 587) to conclude that «inflectional morphology is what is relevant to syntax». I think that this observation is not entirely adequate:

Firstly, the syntactic relevance of inflection implies involvement of such factors as agreement and government. However, Greek inflected forms and their features do not entirely depend on these factors. For instance, the feature of inflection class is an entirely morphologically motivated feature. Its presence is not predicted by any syntactic rule and this feature clearly affects the shape of inflectional material. In a framework that considers inflectional endings to be simple formatives produced by spell - out rules (e.g.,

the Government and Binding theory, Chomsky 1981), such morphological features have to be available all the way through the syntactic component. Such an unnecessary burden is avoided if inflected words are produced within an autonomous morphological component.

Secondly, there may be syntactically distinct features which correspond to single morphological realizations. Take for example, nominative in Greek: there is nominative assigned by an INFL node and nominative assigned to attributes within a VP. It is a fact, however, that when the feature [nominative] is morphologically assigned to an item through the addition of an inflectional affix, the morphology cannot discern the polyvalent syntactic side of this feature.

Finally, with the exception of prepositions, Greek wordforms of major lexical categories are generally obtained by joining an inflectional ending to a stem. In this language, stems can either be simple or morphologically complex formations, whether they are compounds or derivatives. The possibility offered by a unified treatment of all word formation operations seems to be the most promising way of analysing words.

Thus, I have good reason to believe that inflection may be treated morphologically. As I have already stated, inflection in Greek is based on affixation. This places it very close to derivation which is another word formation process mostly based on affixation. Nevertheless, the two processes have a number of different properties[6]. Besides the fact that inflection is a more productive process than derivation, one of the most significant distinctions has to do with the fact that inflectional affixes appear only once in a word construction and are always the last to be used in a concatenation process. Derivational affixed are, instead, word internal constituents and can be applied recursively to stems; the result of such an application is always a stem. Consider, for example, the word 'aγonistikos' «fighting» and its internal structure containing the morphemes 'aγon-' (the noun base) '-iz-', '-ti-', '-ik-', (the derivational suffixes) and '-os-' (the inflectional suffix):

(4) $[[[[[\text{aγon}]_N - \text{iz}]_V - \text{ti}]_N - \text{ik}]_A - \text{os}]_A$[7]

In this example, 'aγon' - «a fight», 'aγoniz' - «to fight», 'aγonisti' - «fighter», 'aγonistik' - «fighting» (adjective) are all stems. This last operation brings all affixation processes to an end. Thus, inflection and derivation are not exactly the same and may be differentiated but within the same morphological component.

By considering inflection to be treated independently of syntax, I follow the generalised lexicalist hypothesis (Lapointe 1980), according to which no syntactic rule has the power to affect word internal structures. My approach prohibits the appearance of bound morphemes, i.e., stems, as terminal items of syntactic representations. I propose, then, that fully inflected words are inserted into appropriate syntactic representations. In such a framework, syntactic agreement and government ensure the correct matching of constituents to appropriate syntactic structures.

---

6. cf. Scalise (1988) for a detailed account of the differences between inflection and derivation.
7. /z/ becomes /s/ and the /i/ of '-ti' disappears by an application of the appropriate phonological rules.

2. Inflection in EUROTRA

In EUROTRA inflection is not considered to be one of the topics of major interest, mainly for two reasons:

a) The project is machine translation oriented and different paradigmatic forms of the same lexical entry of the source language do not influence the choice of the correct entry in the target language. As a matter of fact, surface inflectional forms do not appear in the «canonical» representations of the rather semantically based level of I(nterface) S(tructure) from which translation is realised into the corresponding IS level of the target language. IS, by definition, must contain language independent representations in order to provide as many as possible isomorphic structures for a «simple» transfer between languages (cf. EUROTRA Reference Manual, v. 6.1).

b) A certain number of EUROTRA languages (e.g. English) have a comparatively poor inflectional system. Therefore, inflection had not been given any significant consideration with respect to other interlingual phenomena qualified to be of a major importance.

Nevertheless, it should not be ignored that although inflectional forms do not appear in IS representations, some inflectional features, like aspect or tense, provide semantic information which can be crucial for contrastive purposes. Furthermore, a number of inflectional features do not only characterise lexical units but phrasal nodes as well. For example, tense is a sentence feature and the feature of number marks all noun phrases.

Inflection, therefore, is no less important to treat with respect to other linguistic phenomena. As inflectional material is bound to morphemes (i.e., stems), it is expected that any treatment of inflection should start from a lower level, where inflected words are analysed and generated. Such a level is the so called EMS (EUROTRA Morphological Structure) which includes all word internal phenomena. The choice of treating inflection at EMS is linguistically justified by the one generalised lexicalist hypothesis (cf. section 1.3) according to which inflected forms constitute a morphological domain and syntax has no effect on them. Therefore, EMS could be viewed as an autonomous morphological component containing word formation processes.

2.1 The EMS level

2.1.1 General remarks

EMS is considered to be the level where all morphological objects are analysed into basic strings which correspond more or less to the linguistic entities of morphemes[8]. As stated in the first section, for Greek, a morpheme can be a stem, an affix, or even a word if this word has no internal structure (e.g., simple words like the adverb 'kato' «down»).

When generation takes place, EMS is also responsible for the synthesis of words from smaller units and from featurised information appearing in tree-representations of previous syntactically motivated levels.

Unification is used as the basic tool for analysing and generating the structure of

---

8. A morpheme is always defined on phonological grounds while a basic unit in EMS (string) is delimited by text considerations.

words. The same formalism is also used for syntactic purposes in subsequent levels of syntactic and syntactico - semantic analysis.

In addition to all this, note that EMS, as it is presented here, does not account for information containing typography and text structure. Such information should be taken into consideration in a previous level concerning text handling phenomena[9].

2.1.2 The EMS dictionary

Basic strings (i.e., morphemes) constitute the lexical entries of the dictionary in EMS. According to the format of all other entries in the ET formalism, they are listed as feature bundles containing information of the following kind:

a) Information about the basic form of the entry, expressed as the attribute - value pair «lu = L».

b) Reference to allomorphic variation represented as «string = S»[10]. In theoretical linguistics, different forms of the same lexical entry are considered to be allomorphs only when there is no phonological explanation for the change of the basic form (cf. Lieber 1980, Ralli 1988). In the ET model there is no place for phonology. All different forms of the same entry must be listed in the dictionary. Only recently there has been an attempt to combine the EMS treatment of words with front - end considerations (cf. N. Jaeger' s paper «The Front End of EUROTRA (EUROTRA Internal Paper). This approach, however, needs a further elaboration, mostly on a linguistic basis, in order to be applied to all EUROTRA languages.

c) Morphosyntactic information denoting the morphological class to which an entry belongs (stem, affix, word, etc.), the grammatical category and different inflectional features such as gender, number, tense, etc. depending on the particular entry[11].

d) A number of diacritic features denoting well - formed combinations of bases with inflectional endings (the 'ic' features, cf. section 1). They constitute a special kind of feature which, when it is assigned to both a stem and an inflectional affix permits the right matching between them.

Note that in the set of diacritics we could also list features such as [learned = yes/no] (cf. Selkirk and Dell 1978) which characterise words borrowed from other languages. This kind of feature may express the peculiar behavior of these words towards inflection. Take for instance the case of Greek neuter nouns such as 'film' «film» and 'taksi' «taxi» which are not inflected according to the common plural/case endings for neuter nouns. This absence of inflection could be lexically taken into consideration if we use a feature [learned = yes], which would prevent the words marked for it to be decomposed in

---

9. For a unified account of text handling and inflection, the reader should refer to N. Jaeger' s paper «The Front End of EUROTRA» (EUROTRA Internal Paper).
10. By allomorphs, I do not mean suppletive forms. Note that suppletive forms must be listed as independent dictionary entries.
11. According to the framework followed by the project, frame information is used in subsequent levels, namely in ERS. However, I believe that in order to have a complete morphological level, frames should not be ignored in EMS morphological representations. As a matter of fact, on several occasions (i.e., nominalisations) the addition of a derivational suffix triggers a change or the erasure of the frame of the base.

smaller parts in analysis, or to be combined with inflectional endings in generation. Level ordering information used to represent the order with which different affixes are added to the bases may also be considered as a kind of diacritic feature. Level ordering reflects not only constituency but also lexical phonological phenomena. We could introduce it under the name of [level = V][12].

Finally, semantic information expressed in terms of semantic relations and selectional features could be included in the lexical entries as well[13].

In (5) I give an example of four EMS entries participating in the structure of the Greek words 'anθropos' «man, nominative, singular» and 'trexun' «they run»:

(5)
l_anθrop =
{mc = stem, cat = n, el_nclass = common,
el_lu = anθrop, string = anθrop, el_gender = masc, ic = n1}.

l_trex =
{mc = stem, cat = v, el_vclass = vb, el_lu = trex,
string = trex, el_diath = act, ic = v1}.

l_os =
{mc = infl, el_lu = os, string = os, el_case = nom, nb = sing, ic = n1}.

l_un =
{mc = infl, el_lu = un, string = un, el_voice = active, el_aspect =
imperfective, el_mstense = pres, person = third, nb = plu, ic = v1}[14].

2.1.3 The structure building rules

A morphological level also contains a number of word formation rules responsible for the analysis and generation of morphologically complex words. Word formation rules in EMS are represented by the framework's machinery of structure building rules which give tree - representations with precedence relations between the daughters and fixed domination relations between the daughters and the mother node. The general linguistic rules analysing and generating inflected words are of the following form:

(6) Word → Stem Infl (for stem-based languages, i.e., Greek)
Word → Word Infl (for word-based languages, i.e., English)

---

12. Level ordering information refers to the framework of lexical phonology and morphology (cf. Kiparsky 1982 and Mohanan 1982). Such an information would be very useful for the computational analysis of derivational morphology.
13. Note that information about semantic relations is closely related to frames. Therefore, it will be included in the feature bundles describing EMS dictionary entries only if frames are taken into consideration as well. Note also that the ET system of semantic features needs to be improved. Consequently, the entries given here as examples do not include any semantic information.
14. The abbreviations of certain features used here are widely adopted in all EUROTRA languages and they stand for: mc = morphological class, lu = lexical unit, etc. (cf. EUROTRA Reference Manual, v. 6.0. and 6.1.). The prefix 'el' added to certain attributes denotes specificity for Greek.

This general pattern is easily translated into the form of structure building rules. Examples of structure building rules concerning Greek noun and verb inflection are given below:

(7)
b_noun =
{mc = word, bar = zero, cat = n, el_nclass = T, el_gender = G, nb = N, el_lu = L
el_case = C} [{mc = stem, cat = n, el_nclass = T, el_gender = G, el_lu = L,
string = S, ic = I},
{mc = infl, nb = N, el_case = c, el_lu = U, string = R, ic = I}].

b_verb =
{mc = word, bar = zero, cat = v, el_vclass = vb, el_voice = V, el_diath
= D, el_aspect = A, el_mstense = T, el_lu = L, person = P, nb = N}
[{mc = stem, cat = v, el_vclass = vb, el_diath = D, el_lu = L, string = S, ic = I},
{mc = infl, el_voice = V, el_aspect = A, el_mstense = T, person = P, nb = N,
el_lu = U, string = R, ic = I}].

Note that constraints on the formation of morphological objects may be expressed in terms of filter rules. Constraints are also expressed by certain features listed in the lexical entries such as diacritics and level ordering information.

With respect to features and feature - passing operations, nominal inflection is simpler than verbal inflection. Verbal inflectional features are numerous and may also belong to distinct inflectional affixes that can be juxtaposed within the same structures. Take for instance the case of the future tense in Romance languages where the person/ number affix is added to an already complex structure containing the basic verbal stem and the morphological mark for future:

(8) French: [[[part]-ir]-ai]
Italian: [[[part]-ir]-o]
etc.

Many examples such as the one mentioned above can be found in Greek:

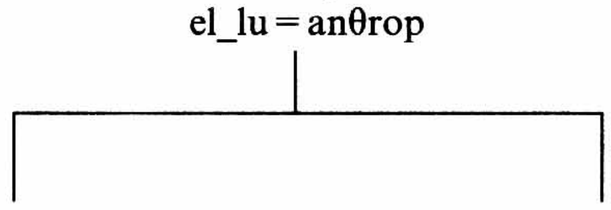
(9) Greek: [[[[teljo]-s]-a]-me]«we finished»
(cf. Babiniotis 1972 and Ralli 1988)

An approach which decomposes verbal inflectional endings into several distinct inflectional affixes is an elegant solution and linguistically motivated. Nevertheless, as it contributes to an increase of computational complexity of the morphological processing, it is perhaps preferable to adopt a treatment according to which the inflectional part of a word is treated as one single affix.

As fas as feature-passing operations are concerned from daughters to mother nodes, it is worth noting that inflectional affixes are not considered to be heads of their constructions (cf. Selkirk 1982 and Ralli 1988). This means that in a morphological tree-representation the feature of grammatical category comes from the stem which assumes the role of the head of the construction. However, inflectional affixes are responsible for giving certain morphosyntactic values to the created inflected words. Take for instance a Greek noun stem listed in the dictionary without any reference to specific morphosyn-

tactic values of case and number. These values are acquired through the combination of the stem with in the appropriate inflectional affix:

10 'anθropos' «man» ([[anθrop]-os])
cat = n[15]
el_nclass = common
el_gender = masc
el_case = nom
nb = sing
el_lu = anθrop



cat = n
el_nclass = common
el_gender = masc
el_lu = anθrop
string = anθrop
ic = n1

el_case = nom
nb = sing
el_lu = os
string = os
ic = n1

As we see in (10), percolation of features to the mother node is not achieved only from the head. Morphosyntactic specifications of inflected words are also due to the non-head which in the case of 'anθropos' is the inflectional affix of nominative singular.

In EMS, the tree - representation proposed above is given by a general structure building rule the formulation of which is already given in (7). One of the features which plays a major role in the formation (or the analysis) of inflectional constructions is the attribute - value pair ([bar = zero]). As I have already said, I consider morphological constructions to be atomic with respect to syntax (cf. section 1). That is, for syntax, the substructure of a morphologically complex object is opaque since the syntactic operations can only inspect the topmost feature bundle. In the ET model, syntactic operations are blocked from applying to the word internal structures with the use of the bar feature ([bar = zero]). This feature characterises the topmost feature bundle and defines it as an atomic unit to the application of operations in subsequent syntactically defined levels of phrasal processing. The postulation of the bar feature is justified by research in theoretical linguistics (Di Sciullo and Williams 1987).

2.2 Inflection in syntax

Inflection is a domain of grammar in which morphological and syntactic considerations overlap to a significant extent. It constitutes an area where morphological properties must be visible by syntactic principles. These properties are reconfirmed by an interaction with the syntactic environments in which they appear. In order to account

---

15. Gender is considered to be a feature of the stem (cf. Ralli 1988 for more details about this).

for the distribution of these morphological properties, inflection must be taken into consideration by areas of syntactic analysis such as agreement and phrasal property realisation. Therefore, a good formalism in theoretical as well as in computational linguistics must explore the kind of information found in the interface between syntax and word structure.

In the ET system, the levels following EMS constitute domains of surface and deep syntactic analysis and generation. In the hierarchical tree - representations of these levels, lexical realisation of inflectional endings is rather redundant. However, information related to inflection is relevant to all grammatical domains other than morphology. As a matter of fact, inflectional features are used for syntactic purposes: in the case of agreement, there is generally a matching between the features of the head and the features of the sister node. A number of inflectional features are also necessary for semantics (e.g., tense, aspect and even definiteness for languages that have it as an inflectional marker). The use, however, of certain inflectional features in levels other than EMS is language specific and it depends on the needs of the grammars.

If we restrict ourselves to inflection in ECS which is the level of surface syntax following (or preceding in case of synthesis) EMS, our job is to represent all morpho-syntactic information contained in the inflectional endings as part of the units that constitute the basic entries for this level's dictionary[16]. In the analysis module, if we arrive to extract information from inflectional endings, relevant to syntax, the formal presence of the affixes themselves is no longer necessary. Then, only stems can assume the role of the dictionary's basic entries. Extraction of appropriate inflectional information as well as erasure of redundant material can be achieved through the passage from EMS to ECS with the help of the translation rule operations. According to these operations, hierarchical representations can be flattened by an erasure of nodes containing redundant material. A t(ranslation) rule performing the needed operation may be the following:

(11) t_word = ˜:{mc = word, bar = zero, el_lu = L}
[˜:{mc = stem, el_lu = L},
˜:{mc = infl}]
⇒ {bar = zero, el_lu = L}.
(The operator « ˜ » denotes deletion of nodes).

Conversely, in generation, our job is to set out inflectional endings as nodes of tree-representations representing inflectional structures. This can be done on the basis of featurised morphosyntactic information found in ECS, already characterised as inflectional, and with the help of a limited number of translation rules. The following translation rule contributes to the creation of word internal structure for Greek nouns:

(12) {cat = n, bar = zero, el_nclass = T, el_lu = L}
⇒ {mc = word, cat = n, bar = zero, el_nclass = T, el_lu = L}
[{mc = stem, el_lu = L, el_nclass = T},
{mc = infl}].

In generation, the intoduction of inflectional material creates the need of new entries

16. In the system, every level has its own dictionary.

in the EMS dictionary where stems and inflectional endings must be listed among the basic entries. For example, the morphematic constituents of the word 'anθropos' will be represented differently in the ECS and EMS dictionaries: the stem 'anθrop' will be the only entry in ECS. Instead, in EMS, 'anθrop' will be listed together with the set of nominal inflectional ending expressing different case/number values, namely, '-os', '-u', etc.. Since inflectional endings are not lexically present in all the syntactically relevant levels (i.e., ECS, ERS and IS) but are only featurised, I shall not refer to the rules of these levels.

2.3 Inflection and machine translation

Let us now examine how the representation of inflection in the ET system is important for translation purposes. ERS and IS make use of a limited number of inflectional features the choice of which is language specific. In analysis, these features are transmitted from ECS. In generation, it is up to the target language to decide whether information received from other languages through transfer could be kept as such or new features have to be added to the already existant features. It may be the case that some inflectional features of the target language have to be ignored and others have to be introduced. For example, definiteness does not constitute a morphologically realised inflectional feature in Greek. Instead, in this language, gender is important not only for agreement matters (e.g., between adjectives and nouns) but also for disambiguation purposes: between two IS entries like 'xor-' «place» (masculine) and 'xor-' «country» (feminine), gender is the main distinguishing feature. Consequently, during transfer between English and Greek, one has to take into consideration the fact that gender will be ignored when English is the target language, but it must be introduced when Greek plays the target role.

An additional argument in favor of stating the importance of inflection in machine translation resides in the fact that inflectional features, contribute to the identification of constituents in a free word order language. For example, in Greek, a subject does not have a fixed position in the sentence and is only identified by the nominative case[17]. Furthermore, the presence of a constituent in genitive case detects a noun modifier being in a dependency relation with its head (the modifier will be in the genitive case) as opposed to a noun modifier marked by nominative case which is in an attributive relation with the head.

3. Conclusions

In this paper, I have offered a description of how inflection in Greek has been taken into consideration by the ET machine translation system. I have considered inflection to be a morphological process and as such I have proposed a treatment similar to the one proposed for other morphological processes, namely derivation and compounding (cf. Eurotra Reference Manual, v. 6.0.). Inflectional endings were given the status of independent dictionary entries, listed as feature bundles. I have shown that inflectional

---

17. In European languages with a poor morphological case system, there is a tendency to substitute the absence of morphological case by an extensive use of prepositions.

features affect the sentential structures into which they appear and determine the range of lexical items that may be inserted in various positions of these structures. Finally, I have shown that although inflection has a minor role to play in the syntactic levels of analysis and generation proves to be of significant importance for disambiguation purposes. Therefore, a computational treatment of inflection is crucial, especially for languages, like Greek, with a rich inflectional system.

Acknowledgments and disclaimer

The research reported on here was undertaken in the framework of the EUROTRA Machine Translation Project, sponsored by the Commission of the European Communities. The views of the author of this paper are not necessarily those of the EUROTRA project management.

Angela Ralli
Dept. of French Studies
University of Athens

BIBLIOGRAPHY

Anderson S. 1982: «Where' s morphology?». *Linguistic Inquiry* 13, 571 - 612.
Anderson S. 1989: «Inflection». In M. Hammond – M. Noonan (eds.), *Theoretical morphology* (Academic Press).
Aronoff M. 1976: *Word formation in Generative Grammar,* Linguistic Inquiry Monograph 1 (MIT Press).
Aronoff M. 1990: *Morphology by itself.* Class lectures, Third European summer school in language, logic and information 1991.
Babiniotis G. 1972: *The verb in Modern Greek* (Athens).
Bauer L. 1988: *Introducing linguistic morphology* (Edinburgh University Press).
Chomsky N. 1981: *Lectures on Government and Binding* (Dordrecht: Foris).
Di Sciullo A.M. – E. Williams 1987: *On the definition of word* (MIT Press).
Ernout A. 1933: *Morphologie historique du Latin* (Paris: Klincksieck. Eurotra Reference manual, versions 6.0. and 6.1).
Jensen J. – M. Stong-Jensen (1984): «Morphology is in the lexicon». *Linguistic Inquiry* 15:3, 474 - 498.
Jensen J. 1990: *Morphology.* Amsterdam Studies in the theory and history of linguistic science (John Benjamins Publ. Co).
Joseph B.D. – I. Philippaki-Warburton 1986: *Modern Greek* (Croom Helm).
Karttunen L. 1983: «KIMMO: A two-level morphological analyser». *Linguistic Forum* 22, 165 - 186.
Kiparsky P. 1982: *Lexical phonology and morphology.* Linguistic Society of Korea, Linguistics in the morning calm (Hanshin, Seoul).
Koskenniemi K. 1983: *Two-level morphology* (Helsinki, Finland: University of Helsinki, Dept. of General Linguistics, Publication No 11).
Lapointe S. 1980: *A theory of grammatical agreement* (Ph. D. diss., University of Amherst).

Lieber R. 1980: *On the organisation of the lexicon* (Ph. D. diss., MIT).
Malikouti-Drachman A. – G. Drachman 1988: «Stress in Modern Greek». *Studies in Greek Linguistics* (Thessaloniki).
Matthews P. 1974: *Morphology* (Cambridge University Press).
Meillet A. 1934: *Introduction à l' étude comparative des langues Indo-Européennes* (Paris).
Mohanan K.P. 1982: *Lexical Phonology* (Ph. D. diss., MIT).
Ralli A. 1988: *Eléments de la morphologie du grec moderne: la structure du verbe* (Ph. D. diss., Université de Montréal).
Ralli A. – E. Galiotou 1987: «A morphological processor for Modern Greek, *ACL Meeting, European Chapter*.
Ralli A. 1986: «Inflection vs. derivation». *Studies in Greek Linguistics* (Thessaloniki).
Ralli A. 1985: *Greek morphology*, Greek volume on the completion of the first phase of the Eurotra project.
Scalise S. 1988: «Inflection and derivation». *Linguistics* 26, 561-581.
Scalise S. 1984: *Generative morphology* (Dordrecht: Foris).
Selkirk E. 1982: *The syntax of words* (MIT Press).
Selkirk E. – F. Dell 1978: *On a morphologically governed alternation*. Linguistic Inquiry, Monograph 3 (MIT Press).
Sproat R. 1992: *Morphology and computation* (MIT Press).
Touratzidis L. – A Ralli 1992: «Stress in Greek inflected forms: a computational treatment». *Language and Speech* 35, 435-453.
Triantaphyllides M. 1939: *Grammar of Modern Greek* (Athens).
Williams E. 1981: «On the notions lexically related and head of the word». *Linguistic Inquiry* 12, 245-274.
Zwicky A. 1985a: «Heads». *Journal of Linguistics* 21, 1-29.